ENTROPY AND ESCAPE OF MASS FOR HILBERT MODULAR SPACES

SHIRALI KADYROV

ABSTRACT. We study the relation between metric entropy and escape of mass for the Hilbert modular spaces with the action of a diagonal element.

1. INTRODUCTION

Many interesting problems are related to equidistribution on homogeneous spaces. Often the ambient space is not compact, which leads to the question whether the limit measure is still a probability measure. One case of this so-called non-escape of mass problem is for a sequence of measures that are invariant under one parameter unipotent subgroups. In this case the answer is simple: for a sequence of invariant and ergodic measures under unipotent subgroups the limit measure is still a probability measure or the zero measure [8]. This fact relies on the quantitative non-divergences estimates for unipotents due to works of S. G. Dani [1] (further refined by G. A. Margulis and D. Kleinbock [7]).

In this paper we are interested in the dynamics of diagonal flows. Consider a sequence of probability measures invariant under a particular diagonal element of a linear group acting on the homogeneous space. In this case, the limit measure of the space could be any value in [0,1]. However, if additionally we assume that the measures have high entropy w.r.t. the diagonal element then one can show that the limit measure is not 0. This has been realized in [3] where M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh show the following.

Theorem 1.1. *Let X be the unit tangent bundle to the modular surface and T be the time 1-map for the geodesic flow. Then, any sequence of T-invariant probability measures μ_n with entropies $h_{\mu_n}(T) \geq c$ satisfies that any weak* limit μ_∞ has at least $\mu_\infty(X) \geq 2c - 1$ mass left.*

Here, μ_∞ is a weak* limit of the sequence $(\mu_n)_{n\geq 1}$ if for some subsequence n_k and for all $f \in C_c(X)$ we have

$$\lim_{k\to\infty} \int_X f d\mu_{n_k} \to \int_X f d\mu_\infty.$$

In [2] M. Einsiedler and the author prove a similar theorem for the space of three-dimensional lattices. Our main goal in this paper is to extend Theorem 1.1 to the following more general setup.

Let F be an algebraic number field and let $\mathcal{O}$ be its ring of integers. Let $S^\infty = \{\sigma_1, ..., \sigma_{r+s}\}$ be its archimedean places where $\{\sigma_1, ..., \sigma_r\}$ are the real places and

the rest are complex ones. Define

$$G := \prod_{n=1}^{r} \mathrm{SL}_2(\mathbb{R}) \times \prod_{m=1}^{s} \mathrm{SL}_2(\mathbb{C}) \text{ and } \Gamma := \mathrm{SL}_2(\mathcal{O}).$$

We have the natural embedding of Γ into G via

$$\Delta : \gamma \to (\sigma_1(\gamma), \sigma_2(\gamma), \dots, \sigma_{r+s}(\gamma))$$

where $\sigma_j(\gamma) = \begin{pmatrix} \sigma_j(a) & \sigma_j(b) \\ \sigma_j(c) & \sigma_j(d) \end{pmatrix}$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$. Then Γ becomes a lattice in G (cf. Lemma 2.2). It is an irreducible lattice and the quotient space $X := \Gamma\backslash G$ is non-compact. The bi-quotient $\Gamma\backslash \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2) \times \mathrm{SO}(2)$ in the case of a real quadratic field F over $\mathbb{Q}$ is known as Hilbert modular surface.

Let a be any fixed diagonal element of G. Then there exist $a_j \in \mathbb{R}$ and $\theta_j \in [0, 2\pi]$ such that

$$a = diag(e^{i\theta_1} e^{a_1/2}, e^{-i\theta_1} e^{-a_1/2}) \times \cdots \times diag(e^{i\theta_{r+s}} e^{a_{r+s}/2}, e^{-i\theta_{r+s}} e^{-a_{r+s}/2})$$

with $\theta_1, \dots, \theta_r = 0$. Now, we define the action of T on X by $\mathrm{T}(x) = x \cdot a$.

In § 2 we define the height function $\mathrm{ht}(\cdot)$ on X. Now, if we define $X_{<M} = \{x \in X : \mathrm{ht}(x) < M\}$ then $X_{<M}$ becomes pre-compact (cf. Lemma 2.2). We similarly define $X_{\geq M}$. Now, we can state the main result.

Let $|a_1| + \cdots + |a_r| = h_r$ and $|a_{r+1}| + \cdots + |a_{r+s}| = h_s$. We note that the maximal metric entropy of T is $h_r + 2h_s$, which we denoted by $h_{\max}(\mathrm{T})$.

Theorem 1.2. *Let $M > \max\{e^{3h_{\max}(\mathrm{T})}, 100\}$ be given. Then, there exists a continuous decreasing function $\phi : \mathbb{R}^+ \to \mathbb{R}$ with $\lim_{M \to \infty} \phi(M) = 0$ such that*

$$\mu(X_{<M}) \geq 1 - \frac{2}{h_{\max}(\mathrm{T})}(h_{\max}(T) - h_\mu(T)) + \phi(M)$$

for any T-invariant probability measure μ on X. In particular, for a sequence of T-invariant measures μ_n with $h_{\mu_n}(\mathrm{T}) \geq h$ one has that any weak* limit μ_∞ has at least $\frac{2h}{h_{\max}(\mathrm{T})} - 1$ mass left.

Whenever $h \in (h_{\max}(\mathrm{T})/2, h_{\max}(\mathrm{T})]$ there will be some mass left in the limit. We think that the theorem is sharp in the following sense: there should exists a sequence of T-invariant probability measures $(\mu_n)_{n \geq 1}$ on X with $\lim_{n \to \infty} h_{\mu_n}(\mathrm{T}) = h_{\max}(\mathrm{T})/2$ such that the limit measure is the 0 measure. A similar construction has been carried out in [6] for the space $\mathrm{SL}_n(\mathbb{Z})\backslash \mathrm{SL}_n(\mathbb{R})$ of unimodular lattices. Theorem 1.2 suggests the following.

Conjecture 1.3. *Let G' be a $\mathbb{Q}$-group and Γ' be an arithmetic lattice of $\mathbb{Q}$-rank one. Let T' be a right multiplication on $\Gamma'\backslash G'$ by a diagonalizable element in G. Then, any sequence of T'-invariant probability measures μ_n on $\Gamma'\backslash G'$ with entropies $h_{\mu_n} \geq c$ satisfies that any weak* limit μ_∞ has at least*

$$\mu_\infty(\Gamma'\backslash G') \geq 1 - \frac{2}{h_{\max}(T')}(h_{\max}(T') - c)$$

mass left where $h_{\max}(T')$ is the maximal metric entropy of T'.

For the $\mathbb{Q}$-rank one case, the conjecture suggests that once the entropies of the measures are uniformly greater than $1/2$ of the maximal entropy, there is always some mass left in the limit. For a heuristic explanation we refer to Remark 5.2 in [3].

Another interesting case studying the limits of a sequence of probability measures arises by averaging an arbitrary measure under iterates of some element of the ambient group. In this case, the notion of entropy does not make sense rather one has to consider the dimension of the measures.

For any group H we define $B_\epsilon^H(g)$ to be the open ball in H of radius $\epsilon > 0$ centered at $g \in H$ and we simply write B_ϵ^H if the ball is centered at the identity 1. Let us consider the following subgroups of G

$$U^+ = \{g \in G : a^{-n}g\, a^n \to 1 \text{ as } n \to -\infty\},$$
$$U^- = \{g \in G : a^{-n}ga^n \to 1 \text{ as } n \to \infty\},$$
$$L = \{g \in G : ga = ag\}.$$

We let $D := \dim U^+ \leq r + 2s$. Let $d \in [0, D]$ be given and let us consider a probability measure ν in X with the following property. For any $\delta > 0$ there exists $\epsilon' > 0$ such that for any $\epsilon < \epsilon'$ one has

$$\nu(xB_\epsilon^{U^+}B_\eta^{U^-L}) \ll \epsilon^{d-\delta} \text{ for any } \eta \in (0,1) \text{ and for any } x \in X.$$

In this case say that ν has a *dimension at least d in the unstable direction*. Now, we consider the following sequence of measures μ_n defined by

$$\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} \mathrm{T}_*^j \nu$$

where $\mathrm{T}_*^j \nu$ is the push-forward of ν under T^j. We have

Theorem 1.4. *For a fixed d let ν be a probability measure of dimension at least d in the unstable direction with respect to a, and let μ_n be as above. Then the sequence of probability measures $(\mu_n)_{n\geq 1}$ satisfies that any weak* limit μ_∞ has at least $\mu_\infty(X) \geq 1 - \frac{2a_*(D-d)}{h_{\max}(\mathrm{T})}$ mass left where $a_* = \max\{|a_i| : i = 1, \ldots, r+s\}$.*

In particular, if ν has full dimension, that is if $d = D$, then the limit μ_∞ is a probability measure. In this case with a minor additional assumption on ν one in fact obtains the equidistribution result, that is, the limit measure μ_∞ is the Haar measure [9]. We say that an element $x \in X$ is *divergent on average with respect to a* if $\lim_{N \to \infty} \frac{1}{N}\{n \in [0, N-1] : \mathrm{T}^n(x) \in K\} = 0$ for any compact set K in X.

We note that if we have a measure ν as above for some d which is supported in the set of points in X that diverge on average then clearly any limit μ_∞ of $(\mu_n)_{\geq 1}$ is the zero measure which implies that $d \leq D - \frac{h_{\max}(\mathrm{T})}{2a_*}$. This hints the following.

Corollary 1.5. *The Hausdorff dimension of the points in X that are divergent on average w.r.t a is at most $\dim G - \frac{h_{\max}(\mathrm{T})}{2a_*}$.*

The proof of the corollary is easily obtained from Theorem 1.4 using [5, Corollary 4.12] and is left to the reader (cf. [2, Corollary 1.7]).

In the next section we will consider some basic facts. In § 3 we state the main ingredients and show how one deduces Theroem 1.2. In § 4 we introduce the partitions and count the number of elements in these partitions. In § 5 we obtain the main proposition and finally, in § 6 we indicate how one proves Theorem 1.4.

useful comments which in particular helped to improve the results of the previous version of the paper.

2. PRELIMINARIES

We consider the space X as a subspace of the space of $\mathcal{O}$-submodules Λ of $(\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ with the following properties:

(i) Λ is an $\mathcal{O}$-submodule generated by two vectors v, w of $(\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$,

(ii) $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ and $w = (w_1', w_1'') \times (w_2', w_2'') \times \cdots \times (w_{r+s}', w_{r+s}'')$ are such that $\det \begin{pmatrix} v_j' & v_j'' \\ w_j' & w_j'' \end{pmatrix} = 1$ for $j = 1, ..., r+s$.

From now on, we use a standard notation $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ for a vector $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$. A similar notation is used for $w \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$. The action of $\mathcal{O}$ on $(\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ is given by $\lambda \cdot v =$

$$(\sigma_1(\lambda)v_1', \sigma_1(\lambda)v_1'') \times (\sigma_2(\lambda)v_2', \sigma_2(\lambda)v_2'') \times \cdots \times (\sigma_{r+s}(\lambda)v_{r+s}', \sigma_{r+s}(\lambda)v_{r+s}'')$$

for any $\lambda \in \mathcal{O}$ and any $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$.

Now, we define the *height* function $\mathrm{ht}(\cdot)$ from X to $\mathbb{R}^+$ as follows. On $\mathbb{R}$ and on $\mathbb{C}$ we consider the usual absolute value $|\cdot|$ and for any (v_j', v_j'') in $\mathbb{R}^2$ or in $\mathbb{C}^2$ by the norm $|\cdot|$ we mean $|(v_j', v_j'')| = \max\{|v_j'|, |v_j''|\}$. For a vector $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ in an $\mathcal{O}$-submodule $\Lambda \in X$ we define the 'norm' by

$$\|v\| = \prod_{j=1}^{r+s} |(v_j', v_j'')|^{\delta_j}$$

where

$$\delta_j = \begin{cases} 1, & \text{if } j \in \{1, 2, \ldots, r\} \\ 2, & \text{if } j \in \{r+1, r+2, \ldots, r+s\} \end{cases}.$$

Now, we define the height of Λ:

$$\mathrm{ht}(\Lambda) := \max\{\|v\|^{-1} : v \in \Lambda - \{0\}\}.$$

We note that this is well defined as $\|v\| \neq 0$ whenever $v \neq 0$.

Definition. A nonzero vector v in an $\mathcal{O}$-submodule Λ is said to be *primitive* if $(Fv) \cap \Lambda = \mathcal{O}v$.

Lemma 2.1. *Up to multiplication by units, for any element $\Lambda \in X$ there can be at most one primitive (short) vector of norm < 1.*

Having only one short vector is crucial throughout the paper. Obtaining similar results as in this paper for spaces that allow more than one primitive short vectors requires different techniques (cf. [2]).

Proof. Assume by contradiction that there are two distinct primitive vectors $e, f \in \Lambda$ such that $\|e\| < 1, \|f\| < 1$ up to multiplication by units. Let $e = (e_1', e_1'') \times (e_2', e_2'') \times \cdots \times (e_{r+s}', e_{r+s}'')$ and $f = (f_1', f_1'') \times (f_2', f_2'') \times \cdots \times (f_{r+s}', f_{r+s}'')$. We pick $v = (v_1', v_1'') \times (v_2', v_2'') \times \cdots \times (v_{r+s}', v_{r+s}'')$ and $w = (w_1', w_1'') \times (w_2', w_2'') \times \cdots \times (w_{r+s}', w_{r+s}'')$ which generate Λ over $\mathcal{O}$ as a submodule and satisfy the property (ii). There are $\lambda_1, \lambda_2, \nu_1, \nu_2 \in \mathcal{O}$ such that $e = \lambda_1 v + \lambda_2 w$ and $f = \nu_1 v + \nu_2 w$. We have

$$\prod_{j=1}^{r+s} \det \begin{pmatrix} e_j' & e_j'' \\ f_j' & f_j'' \end{pmatrix}^{\delta_j} = \prod_{j=1}^{r+s} \det \left(\begin{pmatrix} \sigma_j(\lambda_1) & \sigma_j(\lambda_2) \\ \sigma_j(\nu_1) & \sigma_j(\nu_2) \end{pmatrix} \begin{pmatrix} v_j' & v_j'' \\ w_j' & w_j'' \end{pmatrix} \right)^{\delta_j}$$

where $\{\sigma_1, ..., \sigma_{r+s}\} = S^\infty$. Since, $\det \begin{pmatrix} v'_j & v''_j \\ w'_j & w''_j \end{pmatrix} = 1$ for $j = 1, 2, \ldots, r + s$ we must have

$$\prod_{j=1}^{r+s} (e'_j f''_j - e''_j f'_j)^{\delta_j} = \prod_{j=1}^{r+s} \sigma_j (\lambda_2 \nu_1 - \lambda_1 \nu_2)^{\delta_j}.$$

We now claim that $\lambda_2 \nu_1 \neq \lambda_1 \nu_2$. Otherwise, we see that $\frac{\nu_1}{\lambda_1} e = \nu_1 v + \frac{\nu_1 \lambda_2}{\lambda_1} w = f$ where without loss of generality we assumed that $\lambda_1 \neq 0$. Then, we have $f \in (Fe) \cap \Lambda = \mathcal{O}e$ and $e \in (Ff) \cap \Lambda = \mathcal{O}f$ which imply, upto multiplication by units, that e and f are the same which is a contradiction.

Since $\lambda_1, \lambda_2, \nu_1, \nu_2 \in \mathcal{O}$, from the above claim we obtain that

$$\prod_{j=1}^{r+s} \sigma_j (\lambda_2 \nu_1 - \lambda_1 \nu_2)^{\delta_j} = N(\lambda_2 \nu_1 - \lambda_1 \nu_2) \geq 1$$

where $N(\cdot)$ is the number theoretic norm. It follows that

$$(2.1) \qquad \qquad \prod_{j=1}^{r+s} (e'_j f''_j - e''_j f'_j)^{\delta_j} \geq 1.$$

From Cauchy-Schwartz inequality we see that

$$|(e'_j f''_j - e''_j f'_j)| \leq |(e'_j, e''_j)| \cdot |(f'_j, f''_j)|.$$

Hence,

$$|\prod_{j=1}^{r+s} (e'_j f''_j - e''_j f'_j)^{\delta_j}| \leq \prod_{j=1}^{r+s} |(e'_j, e''_j)^{\delta_j}| \cdot \prod_{j=1}^{r+s} |(f'_j, f''_j)^{\delta_j}| = \|e\| \|f\| < 1.$$

Thus, we obtain a contradiction to (2.1). Therefore, up to multiplication by units, there can be at most one primitive short vector of norm < 1. $\qquad \square$

We will need the following well known fact (see for example [10]).

Lemma 2.2. Γ *is a lattice in* G *and* $X_{<M}$ *is pre-compact.*

The idea of the proof is to embed G as a $\mathbb{Q}$-group in $\mathrm{SL}_{2(r+2s)}(\mathbb{R})$. This identification goes deeper namely that the points of the module are identified with the points of the lattice and the $\|\cdot\|$ function we considered above is just the Euclidean norm on $\mathbb{R}^{2(r+2s)}$. This gives that Γ is a lattice in G and moreover using Mahler's compactness criterion we obtain that $X_{<M}$ is pre-compact.

3. Main Ingredients and the Proof of Theorem 1.2

In this section we will state Lemma 3.2 and Proposition 3.3 without proofs and show how they can be used to deduce Theorem 1.2. To make use of both Lemma 3.2 and Proposition 3.3 we need the following lemma which gives an upper bound for entropy in terms of covers by Bowen balls.

Define a (forward) Bowen N-ball (of radius η) to be the translate xB_N for some $x \in X$ of

$$B_N = \bigcap_{n=0}^{N-1} a^n B_\eta^G a^{-n}$$

where $\eta > 0$ is fixed such that the log map from B_η^G to the Lie algebra of G is injective.

Lemma 3.1. *Let μ be a T-invariant ergodic probability measure on X. For any $N \geq 1$ and $\epsilon > 0$ let $BC(N, \epsilon)$ be the minimal number of (forward) Bowen N-balls needed to cover any particular subset of X of measure bigger than ϵ. Then*

$$h_\mu(\mathrm{T}) \leq \liminf_{N \to \infty} \frac{\log BC(N, \epsilon)}{N}.$$

To prove the lemma one roughly uses the trivial entropy bound, namely

$$H(\xi) \leq \log |\xi|$$

where $|\xi|$ is the number of elements of the partition ξ and the existence of fine partitions with thin boundary. The proof is left to the reader which is very similar to [3, Lemma B.2].

For $M, N \geq 1$ given we define a partition

$$Q_{M,N} := \bigvee_{n=0}^{N-1} \mathrm{T}^{-n}\{X_{<M}, X_{\geq M}\}.$$

Lemma 3.2. *The partition $Q_{M,N}$ has $\ll e^{O(\frac{\log \log M}{\log M})N}$ elements for any $M \geq e^{h_{\max}(\mathrm{T})}$ and $N \in \mathbb{N}$ where implied constants do not depend on M, N.*

Here $X \ll Z$ means that there exists a positive constant c such that $X \leq cZ$. Also, $X \ll_d Z$ means that the constant c depends on d. The proof of Lemma 3.2 is given in § 4.1.

The partition elements of $Q_{M,N}$ can be described by the suitable subsets of $[0, N-1]$ in the sense that for any $Q \in Q_{M,N}$ there exists $\mathcal{V} \subset [0, N-1]$ with integer end points such that

$$Q = \{x \in X : \forall n \in [0, N-1], \mathrm{T}^n(x) \in X_{\geq M} \text{ if and only if } n \in \mathcal{V}\}.$$

In this case, we denote Q by $Q(\mathcal{V})$.

Proposition 3.3. *For any $M > e^{3h_{\max}(\mathrm{T})}$ the partition element $Q(\mathcal{V}) \in Q_{M,N}$ with $Q(\mathcal{V}) \subset X_{<M}$ can be covered by*

$$\ll_M e^{O(\frac{\log \log M}{\log M})N} e^{h_{\max}(\mathrm{T})(N - \frac{|\mathcal{V}|}{2})}$$

Bowen N-balls for any $N \in \mathbb{N}$ where the implied constant in $O(\cdot)$ is independent of M, N.

The proof of Proposition 3.3 easily follows from Proposition 5.1 together with Lemma 4.1 and it is given after the statement of Proposition 5.1.

Proof of the Theorem 1.2. Note first that it suffices to consider ergodic measures. For if μ is not ergodic, we can write μ as an integral of its ergodic components $\mu = \int \mu_t d\tau(t)$ for some probability space (E, τ), see for example [4, Theorem 6.2]. Therefore, we have $\mu(X_{\geq M}) = \int \mu_t(X_{\geq M}) d\tau(t)$, but also $h_\mu(\mathrm{T}) = \int h_{\mu_t}(\mathrm{T}) d\tau(t)$, see for example [11, Thm. 8.4], so that desired estimate follows from the ergodic case.

Suppose that μ is ergodic. Let $M > e^{3h_{\max}(\mathrm{T})}$ be such that $\mu(X_{<M}) > 0$. Later in the proof we will show how one may choose M independent of μ which is crucial in obtaining the last part of the theorem. We would like to apply Lemma 3.1. For this we need to find an upper bound for covering a subset of X of measure ϵ by

Bowen N-balls. Let us fix $\epsilon > 0$ such that $\mu(X_{<M}) > 2\epsilon$. The pointwise ergodic theorem implies

$$\frac{1}{N} \sum_{n=0}^{N-1} 1_{X_{\geq M}}(\mathrm{T}^n(x)) \to \mu(X_{\geq M})$$

as $N \to \infty$ for a.e. $x \in X$. Thus, there is N_0 such that for $N > N_0$ the average on the left will be bigger than $\mu(X_{\geq M}) - \epsilon$ for any $x \in X_1$ for some $X_1 \subset X$ with measure $\mu(X_1) > 1 - \epsilon$. Clearly, for any $N > N_0$ we have $\mu(Z) > \epsilon$ where

$$Z = X_1 \cap X_{<M}.$$

Now, we would like to find an upper bound for the number of Bowen N-balls needed to cover the set Z. Here $N \to \infty$ while ϵ is fixed. We now split Z into the sets $P(\mathcal{V})$ as in Proposition 3.3. By Lemma 3.2 we know that we need $\ll_M e^{O(\frac{\log \log M}{\log M})N}$ many of these. Moreover, by our assumption on X_1 we only need to look at sets $\mathcal{V} \subset [0, N-1]$ with $|\mathcal{V}| \geq (\mu(X_{\geq M}) - \epsilon)N$. On the other hand, Proposition 3.3 gives that each of those sets $Q(\mathcal{V})$ can be covered by $\ll_M e^{O(\frac{\log \log M}{\log M})N} e^{h_{\max}(\mathrm{T})(N - \frac{1}{2}|\mathcal{V}|)}$ Bowen N-balls. Together we see that Z can be covered by

$$\ll_M e^{O(\frac{\log \log M}{\log M})N} e^{h_{\max}(\mathrm{T})(N - \frac{1}{2}|\mathcal{V}|)}$$

Bowen N-balls. Applying Lemma 3.1 we arrive at

$$\begin{aligned} h_\mu(\mathrm{T}) &\leq \liminf_{N \to \infty} \frac{\log BC(N, \epsilon)}{N} \\ &\leq h_{\max}(\mathrm{T})\left(1 - \frac{(\mu(X_{\geq M}) - \epsilon)}{2}\right) + O\left(\frac{\log \log M}{\log M}\right). \end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we get that

$$(3.1) \qquad h_\mu(\mathrm{T}) \leq h_{\max}(\mathrm{T})\left(1 - \frac{\mu(X_{\geq M})}{2}\right) + O\left(\frac{\log \log M}{\log M}\right)$$

which can be rewritten as

$$\mu(X_{<M}) \geq 1 - \frac{2}{h_{\max}(T)}(h_{\max}(T) - h_\mu(T)) + \phi(M)$$

where $\phi(M) = O\left(\frac{\log \log M}{\log M}\right)$.

In the next we will show that the theorem holds for any $M > M_0 := \max\{e^{3h_{\max}(\mathrm{T})}, 100\}$. Clearly the theorem holds if $\mu(X_{<M_0}) > 0$ so that we may assume $\mu(X_{<M_0}) = 0$. Let us define the number M_μ by

$$M_\mu := \inf\{M > M_0 : \mu(X_{<M}) > 0\}.$$

The above argument implies that (3.1) holds for any $M > M_\mu$. If $\mu(X_{<M_\mu}) > 0$ then (3.1) also holds for $M = M_\mu$. Otherwise if $\mu(X_{<M_\mu}) = 0$ then

$$\lim_{n \to \infty} \mu(X_{\geq M_\mu + \frac{1}{n}}) = \mu(X_{>M_\mu}) = \mu(X_{\geq M_\mu}) = 1.$$

Now, using (3.1) for $M + 1/n$ instead of M and taking the limit as $n \to \infty$ we get (3.1) for $M = M_\mu$. For any $M \in [M_0, M_\mu)$ we need to prove that (3.1) holds. Since $\mu(X_{\geq M}) = 1$ for $M \leq M_\mu$ we see that (3.1) simplifies to

$$h_\mu(\mathrm{T}) \leq \frac{h_{\max}(\mathrm{T})}{2} + O\left(\frac{\log \log M}{\log M}\right).$$

Since $\frac{\log\log M}{\log M}$ is decreasing for $M \geq 100$ and since the above equation holds for $M = M_\mu$ it clearly holds for any $M \in [M_0, M_\mu)$.

For any $M > \max\{e^{3h_{\max}(\mathrm{T})}, 100\}$, one can approximate the characteristic function of $X_{<M}$ by continuous functions with compact support and use (3.1) to obtain the last part of the theorem. $\square$

4. Partitions

For given $M, N \geq 1$ we recall the partition $Q_{M,N} := \bigvee_{n=0}^{N-1} \mathrm{T}^{-n}\{X_{<M}, X_{\geq M}\}$. In this section we estimate the upper bound for the cardinality of $Q_{M,N}$ to prove Lemma 3.2. Later we consider the refinement $P_{M,N}$ of the original partition $Q_{M,N}$ which is crucial in obtaining Proposition 5.1.

From now on, for simplicity, we assume that $a_j \geq 0$ for any $j \in [1, r+s]$. This in particular implies that the unstable subgroup U^+ is a subgroup of *lower* unipotent matrices in G. Also, with this assumption a component vector (v'_j, v''_j) under the iterations of T is getting short, that is $|(v'_j, v''_j)| > |(v'_j e^{a_j/2}, v''_j e^{-a_j/2})|$, means that $|v'_j| > |v''_j| e^{a_j/2}$ as otherwise if $a_j < 0$ then we would get $|v''_j| > |v'_j| e^{a_j/2}$. Hence, the assumption $a_j \geq 0$ is simply a matter of ordering the coordinates of component vectors.

4.1. **Proof of Lemma 3.2.** For any x, the partition element of $Q_{M,N}$ containing x describes the time moments in $[0, N-1]$ for which x stays above height M (and hence when it is below height M) under the action of T. So, we need to calculate the possible configurations of times in $[0, N-1]$. Our main tool to calculate the upper bound for the possible configurations is Lemma 2.1. If there is a time when a point x (under the action of T) is above height M then there is a considerable gap until the next time (if any) when x reaches height M again. This is because the vectors in x can get short (under the action of T) at most once and for another vector in x to become short the earlier vector has to become of norm 1 at least. Now, we explicate the above discussion. Assume that for a vector $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ we have $\|v\| = \prod_{j=1}^{r+s} |(v'_j, v''_j)|^{\delta_j} > 1$. We would like to know an estimate for the smallest possible time n for which the vector v reaches the norm $\leq 1/M$ under the action of T. It is easy to see that the best possible n occurs for example when $v'_j = 0$ for $j = 1, ..., r+s$. In this case, at time n we must have

$$\prod_{j=1}^{r+s} |(0, v''_j e^{-(i\theta_j + a_j/2)n})|^{\delta_j} \leq \frac{1}{M}.$$

Since $\|v\| > 1$, we must have $e^{-\frac{n}{2} \sum_{j=1}^{r+s} a_j \delta_j} < 1/M$ which gives

$$n \geq \frac{2\log M}{h_r + 2h_s}.$$

Similarly, for a vector of norm at most $1/M$, under the action of T, the smallest possible time moment when the norm becomes greater than 1 is again $\geq \frac{2\log M}{h_r + 2h_s}$. We also note that for any vector v in x if the sequence $(\|\mathrm{T}^n(v)\|)_{n \geq 0}$ gets increased at some time then it becomes monotone increasing from that time moment. Thus, in a time interval of length $2\lfloor \frac{2\log M}{h_r + 2h_s} \rfloor$, for any point x in X there can be at most one time interval on which x stays above height M. Hence, $Q_{M, \lfloor \frac{2\log M}{h_r + 2h_s} \rfloor}$ has at most

$$\binom{2\lfloor \frac{2\log M}{h_r+2h_s}\rfloor}{2} \ll \log^2 M \text{ many elements. On the other hand, to obtain } Q_{M,N} \text{ we}$$

need to take refinements of $\lfloor \frac{N}{2\lfloor \frac{2\log M}{h_r+2h_s}\rfloor-1}\rfloor$ many pre-images of $Q_{M,\lfloor \frac{2\log M}{h_r+2h_s}\rfloor}$ and at

most $2\lfloor \frac{2\log M}{h_r+2h_s}\rfloor - 1$ many of $\{X_{<M}, X_{\geq M}\}$. For $M \geq e^{h_r+2h_s}$ we have

$$\left\lfloor \frac{N}{2\lfloor \frac{2\log M}{h_r+2h_s}\rfloor-1}\right\rfloor < \frac{N}{\frac{4\log M}{h_r+2h_s}-3} \leq \frac{N(h_r+2h_s)}{\log M}$$

Hence, we obtain that the cardinality of $Q_{M,N}$ is

$$\ll (\log^2 M)^{\frac{N(h_r+2h_s)}{\log M}} \leq e^{\frac{2(h_r+2h_s)\log\log M}{\log M}N}.$$

$\square$

4.2. The refined partition $P_{M,N}$.

We now consider the refinement $P_{M,N}$ of $Q_{M,N}$. It is a bit technical and the reason why this refinement is needed comes from the product structure of the space G and in particular the way we define the height function $\mathrm{ht}(\cdot)$. The partition elements of $Q_{M,N}$ give information as when the trajectory of a point under T goes into the cusp and when it comes back. Due to the way the height $\mathrm{ht}(\cdot)$ is defined this does not provide much information on individual components (v'_j, v''_j) of the short vectors $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ even if we know that $\|\mathrm{T}^n v\|$ is decreasing on some time interval in $[0, N-1]$. Thus, what we really need is a partitioning of the space X which describes whether components of short vectors under iterates of T decreases or increases. On the other hand, if a component vector gets shorter in n iterates under T, that is, if $|\mathrm{T}^n(v'_j, v''_j)| = |(v'_j e^{na_j/2}, v''_j e^{-na_j/2})| \leq |(v'_j, v''_j)|$, then it is easy to see that we must have $|v'_j|e^{na_j/2} \leq |v''_j|$. This simple observation hints the importance of knowing the ratios $\frac{|v'_j|}{|v''_j|}$ of component vectors. Thus, elements of our new partition $P_{M,N}$ should describe these ratios (cf. (4.6)) of component vectors of short vectors as we define now.

Our goal is to refine the partition $Q_{M,N}$ further by partitioning most of its elements. Let Q be one of its elements. Then there exists $\mathcal{V} \subset [0, N-1]$ such that

(4.1) $\quad Q := Q(\mathcal{V})$
$$= \{x \in X : \text{ for all } n \in [0, N-1], \mathrm{T}^n(x) \in X_{\geq M} \text{ if and only if } n \in \mathcal{V}\}.$$

We split $\mathcal{V}$ into maximal intervals $V_1^{\mathcal{V}}, \ldots, V_k^{\mathcal{V}}$ for some $k \in \mathbb{N}$. For $m = 1, 2, \ldots, k$ we write $V_m^{\mathcal{V}} = [b_m^{\mathcal{V}}, b_m^{\mathcal{V}} + \ell_m^{\mathcal{V}}]$.

For any $j \in [1, r+s]$ recall the fixed number a_j appeared in the definition of T. For each $j \in [1, r+s]$ and $m \in \{1, 2, \ldots, k\}$ let us decompose the extended reals into the following $\ell_m^{\mathcal{V}} + 2$ subintervals:

(4.2) $\qquad I_{0,j}(V_m^{\mathcal{V}}) = [-\infty, b_m^{\mathcal{V}}], \ I_{\ell_m^{\mathcal{V}}+1,j}(V_m^{\mathcal{V}}) = (b_m^{\mathcal{V}} + \ell_m^{\mathcal{V}} a_j, \infty],$

(4.3) $\qquad I_{n,j}(V_m^{\mathcal{V}}) = (b_m^{\mathcal{V}} + (n-1)a_j, b_m^{\mathcal{V}} + na_j] \text{ for } n \in [1, \ell_m^{\mathcal{V}}].$

We write

$$\mathcal{I}_j(V_m^{\mathcal{V}}) = \{I_{n,j}(V_m^{\mathcal{V}}) : n \in [0, \ell_m^{\mathcal{V}}+1]\} \text{ for } m \in \{1, 2, \ldots, k\} \text{ and } j \in [1, r+s].$$

We first note that for any $x \in Q$ there exists a unique primitive vector $v \in \mathrm{T}^{b_m^{\mathcal{V}}-1}(x)$ such that

$$(4.4) \qquad \qquad \| \mathrm{T}^n(v) \| \leq \frac{1}{M} \text{ for } n \in [1, \ell_m^{\mathcal{V}} + 1].$$

We fix $m \in \{1, 2, \ldots, k\}$ and for each $j \in [1, r+s]$ we pick one interval $J_j(V_m^{\mathcal{V}})$ from the set $\mathcal{I}_j(V_m^{\mathcal{V}})$ and consider the product set

$$(4.5) \qquad \qquad J(V_m^{\mathcal{V}}) = J_1(V_m^{\mathcal{V}}) \times \cdots \times J_{r+s}(V_m^{\mathcal{V}}).$$

Now, for any such product set $J(V_m^{\mathcal{V}})$ we associate a partition element, which could be empty, in Q given by

$$(4.6) \quad Q(J(V_m^{\mathcal{V}})) := \{ x \in Q : \exists v \in \mathrm{T}^{b_m^{\mathcal{V}}-1}(x) \text{ such that (4.4) holds and}$$
$$|v_j''| = |v_j'| e^{s_j} \text{ for some } s_j \in J_j(V_m^{\mathcal{V}}) - b_m^{\mathcal{V}} \}.$$

For any $m \in \{1, 2, \ldots, k\}$ we fix one partition element $Q(J(V_m^{\mathcal{V}}))$ as in (4.6) and define the following further refined partition element

$$(4.7) \qquad \qquad P(\mathcal{V}) = \bigcap_{m=1}^{k} Q(J(V_m^{\mathcal{V}})).$$

In this way, for any choice of $Q \in Q_{M,N}$ and any choice of $J(V_m^{\mathcal{V}})$ as in (4.5) we obtain one partition element which is contained in Q. The collection of all possible $P(\mathcal{V})$ as in (4.7) gives a refined partition $P_{M,N}$ of $Q_{M,N}$.

For further motivation why the partition $P_{M,N}$ is crucial we refer to § 5.1, in particular see Lemma 5.2.

Lemma 4.1. *For $M > e^{3h_{\max}(\mathrm{T})}$ and $N \in \mathbb{N}$ the cardinality of the partition $P_{M,N}$ constructed above is $\ll e^{O(\frac{\log \log M}{\log M})N}$ where the implied constants are independent of M, N.*

Proof. Consider a partition element $Q(\mathcal{V})$ of $Q_{M,N}$ as in (4.1). Let $Q(J(V_m^{\mathcal{V}}))$ be as in (4.6) and $P(\mathcal{V})$ be as in (4.7). There are at most $(|V_m^{\mathcal{V}}| + 2)^{r+s}$ possible ways to choose $J(V_m^{\mathcal{V}})$ and hence $(|V_m^{\mathcal{V}}| + 2)^{r+s}$ possible ways to choose $Q(J(V_m^{\mathcal{V}}))$ for a fixed $m \in [1, k]$. Thus, the number of partition elements of $P_{M,N}$ contained in $Q(\mathcal{V})$ is

$$(|V_1^{\mathcal{V}}| + 2)^{r+s} (|V_2^{\mathcal{V}}| + 2)^{r+s} \cdots (|V_k^{\mathcal{V}}| + 2)^{r+s}$$
$$= \exp\left((r+s)[\log(|V_1^{\mathcal{V}}| + 2) + \log(|V_2^{\mathcal{V}}| + 2) + \cdots + \log(|V_k^{\mathcal{V}}| + 2)] \right).$$

This is

$$\ll \exp((r+s) \log(|V_1^{\mathcal{V}}||V_2^{\mathcal{V}}|...|V_k^{\mathcal{V}}|)).$$

We have

$$|V_1^{\mathcal{V}}||V_2^{\mathcal{V}}|...|V_k^{\mathcal{V}}| \leq \left(\frac{|V_1^{\mathcal{V}}| + |V_2^{\mathcal{V}}| + ... + |V_k^{\mathcal{V}}|}{k} \right)^k \leq \left(\frac{N}{k} \right)^k.$$

Also, note that for the function $f(x) = (\frac{N}{x})^x = (N)^x e^{-x \log x}$ its derivative

$$f'(x) = (N)^x \log(N) e^{-x \log x} + (N)^x e^{-x \log x}(-\log x - 1)$$
$$= (N)^x e^{-x \log x}(\log(N) - \log x - 1).$$

Hence $f(x) = (\frac{N}{x})^x$ is increasing on $[1, \frac{N}{e}]$. On the other hand, from the proof of Lemma 3.2 we know that

$$k \leq \left\lceil \frac{N}{2\lfloor \frac{2\log M}{h_r + 2h_s} \rfloor} \right\rceil \leq \left\lceil \frac{(h_r + 2h_s)N}{2\log M} \right\rceil \leq \max\left\{ 1, \frac{(h_r + 2h_s)N}{\log M} \right\}.$$

If $k = 1$ then $(\frac{N}{k})^k = N$. Otherwise, $k \leq \frac{(h_r+2h_s)N}{\log M}$ and for $M \geq e^{e(h_r+2h_s)}$ we have

$$\left(\frac{N}{k}\right)^k \leq \left(\frac{N}{\frac{(h_r+2h_s)N}{\log M}}\right)^{\frac{(h_r+2h_s)N}{\log M}} = \left(\frac{\log M}{h_r + 2h_s}\right)^{\frac{(h_r+2h_s)N}{\log M}}.$$

Hence, the number of partition elements of $P_{M,N}$ contained in $Q(\mathcal{V})$ is $\ll e^{(r+s)\log(N)}$ if $k = 1$ and otherwise it is

$$\ll \exp\left((r+s)\log\left(\left(\frac{\log M}{r+s}\right)^{\frac{(h_r+2h_s)N}{\log M}}\right)\right) \ll e^{O(\frac{\log\log M}{\log M})N}.$$

In either case, the number of partition elements of $P_{M,N}$ contained in $Q(\mathcal{V})$ is $\ll e^{O(\frac{\log\log M}{\log M})N}$. Thus, together with Lemma 3.2 we deduce that $P_{M,N}$ has $\ll e^{O(\frac{\log\log M}{\log M})N}$ elements for $M > e^{3h_{\max}(T)}$. $\qquad\square$

5. MAIN PROPOSITION

In this section we calculate the number of Bowen N-balls needed to cover each partition element of $P_{M,N}$. We recall that a Bowen N-ball is a translate of $B_N = \bigcap_{n=0}^{N-1} a^n B_\eta^G a^{-n}$ in X. We note that the Bowen balls are balls in a different metric that induces the same topology.

Let $M, N \geq 1$ be given. Let $P(\mathcal{V})$ be a partition element of $P_{M,N}$ as in (4.7) such that $P(\mathcal{V}) \subset X_{<M}$. We recall that by definition $\mathcal{V}$ is a subset of $[0, N-1]$ and for all $n \in [0, N-1]$ we have that $T^n(x) \in X_{\geq M}$ if and only if $n \in \mathcal{V}$. In particular, the additional restrictive assumption above is equivalent to $\mathcal{V}$ being in $(0, N-1]$.

Proposition 5.1. *The partition element $P(\mathcal{V}) \in P_{M,N}$ with $P(\mathcal{V}) \subset X_{<M}$ can be covered by*

$$\ll_M c_0^{\frac{h_{\max}(T)}{\log M}N} e^{h_{\max}(T)(N - \frac{|\mathcal{V}|}{2})}$$

Bowen N-balls for some universal constant $c_0 \geq 1$.

Proof of Proposition 3.3. We note that we partitioned any element $Q(\mathcal{V}) \in Q_{M,N}$ into elements $P(\mathcal{V})$ of $P_{M,N}$. From Lemma 4.1 we know that there are at most $\ll e^{O(\frac{\log\log M}{\log M})N}$ such elements of $P_{M,N}$ for $M > e^{3h_{\max}(T)}$. On the other hand, using Proposition 5.1 we deduce that each such element $P(\mathcal{V})$ can be covered by $\ll_M c_0^{\frac{h_{\max}(T)}{\log M}N} e^{h_{\max}(T)(N - \frac{|\mathcal{V}|}{2})}$ Bowen N-balls. By enlarging the implicit constant in $O(\cdot)$ we may assume that $c_0^{\frac{h_{\max}(T)}{\log M}N} \leq e^{O(\frac{\log\log M}{\log M})N}$. Thus, we conclude that any partition element $Q(\mathcal{V})$ with $\mathcal{V} \in (0, N-1]$ can be covered by $\ll_M e^{O(\frac{\log\log M}{\log M})N} e^{h_{\max}(T)(N - \frac{|\mathcal{V}|}{2})}$ Bowen N-balls which completes the proof. $\qquad\square$

We now return to the statement of Proposition 5.1. Roughly, we note that since the number of elements of $P_{M,N}$ is slow exponential as $N \to \infty$, to calculate the entropy it is sufficient to consider the covers of each partition element $P_{M,N}$ by

Bowen balls. Since we only need to count the number of covers of most of the space X (cf. Lemma 3.1) it is reasonable to consider only the partitions $P(\mathcal{V}) \in P_{M,N}$ with $P(\mathcal{V}) \subset X_{<M}$. It is not hard to show that each such partition element $P(\mathcal{V})$ can be covered by $\ll e^{h_{\max}(\mathrm{T})N}$ Bowen N-balls. Thus, the significant factor in Proposition 5.1 is $e^{-\frac{h_{\max}(\mathrm{T})}{2}|\mathcal{V}|}$. Before we start proving Proposition 5.1 we need some preliminary preparations.

5.1. Restrictions of perturbations.
If there are two points in $X_{<M}$ which are η-close to each other such that they both stay above height M for some time interval, then we would like to say that these points must be even closer to each other in the unstable direction U^+. This is not true in general. However, if additionally we know that they are in the same partition element of $P_{M,N}$ then we will show that this is indeed the case.

As before let U^+, U^-, L be the unstable, stable, and centralizer subgroups of G w.r.t. a respectively. We naturally embed U^+ into $\mathbb{R}^r \times \mathbb{C}^s$. We let $u^+(\mathbf{t}) \in U^+$ be the element that corresponds to $\mathbf{t} = (t_1, t_2, \ldots, t_{r+s}) \in \mathbb{R}^r \times \mathbb{C}^s$. For the rest of the section we fix one $P(\mathcal{V}) \in P_{M,N}$ as in (4.7). Recall that $V_m^{\mathcal{V}} = [b_m^{\mathcal{V}}, b_m^{\mathcal{V}} + \ell_m^{\mathcal{V}}], m = 1, 2, \ldots, k$ are the maximal intervals such that $\mathcal{V} = \cup_{m=1}^k V_m^{\mathcal{V}}$. We fix $V_m^{\mathcal{V}}$ for some $m = 1, \ldots, k$ and for simplicity we denote $V_m^{\mathcal{V}} = [b, b+\ell]$. From (4.7) we know that $P(\mathcal{V}) = \bigcap_{m=1}^k Q(J(V_m^{\mathcal{V}}))$ for some $Q(J(V_m^{\mathcal{V}}))$ as in (4.6), namely

$$(5.1) \quad Q(J(V_m^{\mathcal{V}})) := \{x \in Q : \exists v \in \mathrm{T}^{b-1}(x) \text{ such that } (4.4) \text{ holds and}$$
$$|v_j''| = |v_j'|e^{s_j} \text{ for some } s_j \in J_j(V_m^{\mathcal{V}}) - b\}.$$

Lemma 5.2. *Let $x, y \in P(\mathcal{V}) \cap \mathrm{T}^{N-1}(X_{<M})$ with $\mathrm{T}^{b-1}(y) = \mathrm{T}^{b-1}(x)u^+(\mathbf{t})g$ for some $u^+(\mathbf{t}) \in B_{\eta/2}^{U^+}$ and $g \in B_{\eta/2}^{U^- L}$. Then for any $j \in \{1, 2, \ldots, r+s\}$ we have $|t_j| \ll e^{b-n_j}$ where n_j is the left end point of the interval $J_j(V_m^{\mathcal{V}})$.*

Proof. If $J_j(V_m^{\mathcal{V}}) = [-\infty, b] = I_{0,j}(V_m^{\mathcal{V}})$ then $n_j = -\infty$ and in this case the lemma is trivial. So, we may assume $J_j(V_m^{\mathcal{V}}) \neq I_{0,j}(V_m^{\mathcal{V}})$ so that $n_j \geq b$.

By maximality of $V_m^{\mathcal{V}}$ and the fact that $\mathcal{V} \subset (0, N-1]$ we know that

$$\mathrm{T}^{b-1}(x), \mathrm{T}^{b-1}(y) \in X_{<M} \text{ and}$$
$$\mathrm{T}^n(\mathrm{T}^{b-1}(x)), \mathrm{T}^n(\mathrm{T}^{b-1}(y)) \in X_{\geq M} \text{ for any } n \in [1, l+1].$$

Thus there exist vectors $v \in \mathrm{T}^{b-1}(x)$ and $w \in \mathrm{T}^{b-1}(y)$ such that (4.4) holds, that is

$$(5.2) \qquad \qquad \|\mathrm{T}^n(v)\|, \|\mathrm{T}^n(w)\| \leq 1/M \text{ for } n \in [1, \ell+1].$$

On the other hand, from (5.1) for v, w in the standard notation we know that

$$|v_j''| = |v_j'|e^{s_j} \text{ and } |w_j''| = |w_j'|e^{r_j} \text{ for some } s_j, r_j \in J_j(V_m^{\mathcal{V}}) - b.$$

We note that $v_j'' \neq 0 \neq w_j''$ since $(v_j', v_j''), (w_j', w_j'') \neq (0,0)$ (they are rows of matrices of determinant equal to 1) and $s_j, r_j \geq 0$. In particular, if n_j is the left end point of the interval $J_j(V_m^{\mathcal{V}})$ then we have

$$(5.3) \qquad \qquad \frac{|v_j'|}{|v_j''|} \leq e^{b-n_j} \text{ and } \frac{|w_j'|}{|w_j''|} \leq e^{b-n_j}.$$

Also, we know that $w = vu^+(\mathbf{t})g$. So, for $g = (g_1, \ldots, g_{r+s})$ we have $(w'_j, w''_j) = (v'_j, v''_j) \begin{pmatrix} 1 & 0 \\ t_j & 1 \end{pmatrix} g_j = (v'_j + t_j v''_j, v''_j)g_j$ (under the assumption that $a_j \geq 0$ where a_j is as in the definition of a). For $g_j = \begin{pmatrix} d & u \\ 0 & 1/d \end{pmatrix}$ we obtain that

$$(w'_j, w''_j) = (d(v'_j + t_j v''_j), u(v'_j + t_j v''_j) + v''_j/d).$$

Now from (5.3) we get

$$e^{b-n_j} \geq \frac{|w'_j|}{|w''_j|} = \frac{|d(v'_j + t_j v''_j)|}{|u(v'_j + t_j v''_j) + v''_j/d|} \gg \frac{|v'_j + t_j v''_j|}{|v''_j|} = \left| \frac{v'_j}{v''_j} + t_j \right|$$

since d is close to 1 and u is close to 0. Together with (5.3) we deduce that

$$|t_j| \ll e^{b-n_j}.$$

$\square$

Lemma 5.2 alone does not tell us if x, y should be even closer to each other in the unstable direction since for example n_j could be equal to b. Even if $n_j > b$ we still do not know an effective lower bound for n_j. This is because we have only considered one part of the defining properties of $Q(J(V_m^{\mathcal{V}}))$. We have not considered the fact that x, y stay above height M in $[1, \ell + 1]$. In the next lemma we use this fact to obtain the relation among the intervals $J_j(V_m^{\mathcal{V}})$.

Lemma 5.3. *Let $J(V_m^{\mathcal{V}})$ be as in (4.5) and consider $x \in Q(J(V_m^{\mathcal{V}}))$ with $v \in \mathrm{T}^{b-1}(x)$ as in (5.1). Let $S = \{s_1, ..., s_{r+s}\}$ and $\{i_1, ..., i_L\}$ be the subset of S which are ≤ 0, let $j_1, ..., j_C$ be the subset of S such that $s_{j_i} \in (0, (\ell + 1)a_{j_i})$, and let $k_1, ..., k_R$ be the subset of S such that $s_{k_i} > (\ell + 1)a_{k_i}$. In particular, $L + C + R = r + s$. Then*

$$(\ell + 1) \left(\sum_{n=1}^{L} (a_{i_n} k_{i_n}) + \sum_{n=1}^{C} (a_{j_n} k_{j_n}) - \sum_{n=1}^{R} (a_{k_n} k_{k_n}) \right) < 2 \sum_{n=1}^{C} (s_{j_n} k_{j_n}).$$

Proof. Let us consider the j-th component vector (v'_j, v''_j) of v. T acts on v and hence it acts on each of its components and we have

$$\mathrm{T}^n((v'_j, v''_j)) = (v'_j e^{in\theta_j} e^{na_j/2}, v''_j e^{-in\theta_j} e^{-na_j/2})$$

where as before $\theta_j = 0$ if $j \leq r$, and $a_j \geq 0$ for any $j \in [1, r + s]$. Thus,

$$|\mathrm{T}^n((v'_j, v''_j))| = \max\{|v'_j e^{na_j/2}|, |v''_j e^{-na_j/2}|\} = \begin{cases} |v''_j| e^{-na_j/2} & \text{if } na_j < s_j \\ |v'_j| e^{na_j/2} & \text{if } na_j \geq s_j \end{cases}$$

since $|v'_j| e^{s_j/2} = |v''_j| e^{-s_j/2}$. We also note that

$$|(v'_j, v''_j)| = \begin{cases} |v'_j| & \text{if } s_j \leq 0 \\ |v''_j| & \text{if } s_j > 0 \end{cases}.$$

Together we get

$$(5.4) \qquad \frac{|\mathrm{T}^{\ell+1}((v'_j, v''_j))|}{|(v'_j, v''_j)|} = \begin{cases} e^{\frac{(\ell+1)a_j}{2}} & \text{if } s_j \leq 0 \\ e^{\frac{(\ell+1)a_j}{2} - s_j} & \text{if } s_j \in (0, (\ell+1)a_j] \\ e^{-\frac{(\ell+1)a_j}{2}} & \text{if } s_j > (\ell+1)a_j. \end{cases}$$

By the assumption (4.4) on the vector $v \in (\mathbb{R}^2)^r \times (\mathbb{C}^2)^s$ we have

$$\|v\| > \frac{1}{M} \text{ and } \|\operatorname{T}^n(v)\| \leq \frac{1}{M} \text{ for } n \in [1, \ell+1].$$

In particular, this gives

(5.5) $$\frac{\|\operatorname{T}^{\ell+1}(v)\|}{\|v\|} < 1.$$

Now, from (5.4) and (5.5) we get

$$\frac{\prod_{j=1}^{r+s} |\operatorname{T}^{\ell+1}((v_j', v_j''))|^{\delta_j}}{\prod_{j=1}^{r+s} |(v_j', v_j'')|^{\delta_j}}$$

$$= \exp\left(\frac{\ell+1}{2} \sum_{n=1}^{L} (a_{i_n} k_{i_n}) + \frac{\ell+1}{2} \sum_{n=1}^{C} (a_{j_n} k_{j_n}) - \sum_{n=1}^{C} (s_{j_n} k_{j_n}) - \frac{\ell+1}{2} \sum_{n=1}^{R} (a_{k_n} k_{k_n}) \right) < 1.$$

The exponent simplifies to

$$(\ell+1) \left(\sum_{n=1}^{L} (a_{i_n} k_{i_n}) + \sum_{n=1}^{C} (a_{j_n} k_{j_n}) - \sum_{n=1}^{R} (a_{k_n} k_{k_n}) \right) < 2 \sum_{n=1}^{C} (s_{j_n} k_{j_n}).$$

$\square$

The next lemma shows how we apply the above two lemmas. The reader can skip the lemma and come back when it is mentioned in the proof of Proposition 5.1. Recall the embedding of U^+ into $\mathbb{R}^r \times \mathbb{C}^s$.

Lemma 5.4. *Let $V_m^{\mathcal{V}} = [b, b+\ell]$ and $Q(J(V_m^{\mathcal{V}}))$ be as before and let B', B'' be given positive constants. Let us consider the set $D := \{u(\mathbf{t}) \in U^+ : |t_j| < B' \min\{\eta, e^{b-n_j}\}, j = 1, \ldots, r+s\}$ where n_j is the left end point of the interval $J_j(V_m^{\mathcal{V}})$. Then the set D can be decomposed into*

$$\ll e^{\frac{h_{\max}(\mathrm{T})}{2}\ell}$$

disjoint sets of the form $E := \{u(\mathbf{t}) \in U^+ : |t_j| < B'' \eta e^{-\ell a_j}, j = 1, \ldots, r+s\}$.

It is easy to see that the set E is roughly the unstable part of a Bowen ℓ-ball which hints the relation to Proposition 5.1. One can see that a unit ball in U^+ can be covered by $\ll e^{h_{\max}(\mathrm{T})\ell}$ translates of the set E. However, from the lemma above we see that if we consider a subset of a unit ball in U^+ whose elements stay above height M under T on $[b, b+\ell]$ and moreover if the elements of this subset behave similarly, that is, if the set is in the partition $Q(J(V_m^{\mathcal{V}}))$ then we have save $h_{\max}(\mathrm{T})\ell/2$ in the exponent.

Proof. For any $j = 1, 2, \ldots, r+s$ let us consider the ball around 0 of radius $B' \min\{\eta, e^{b-n_j}\}$ in $\mathbb{R}$ or in $\mathbb{C}$ depending whether $j \leq r$ or not and decompose it into the small balls of radius $B'' \eta e^{-\ell}$. If $n_j < b$ (in which case $n_j = -\infty$) then there are $\ll e^{\ell a_j}$ small subintervals if $j \leq r$ and there are $\ll e^{2\ell a_j}$ small balls if $j > r$. Suppose $n_j \geq b$. If $j \leq r$ then there are $\ll e^{\ell a_j + b - n_j}$ small subintervals and if $j > r$ then there are $\ll e^{2(\ell a_j + b - n_j)}$ small balls. We note that if $n_j \geq b + \ell a_j$ (in which case $n_j = b + \ell a_j$) then there are $\ll 1$ small subintervals or $\ll 1$ small balls depending on j. We have $i_1, \ldots, i_L, j_1, \ldots, j_C, k_1, \ldots, k_R$ as in Lemma 5.3. Now, let $i_1', \ldots, i_{L'}'$ be the subset of $\{i_1, \ldots, i_L\}$ which are $\leq r$ and $i_1'', \ldots, i_{L''}''$ be the rest. Similarly, we consider the subsets $j_1', \ldots, j_{C'}'$ and $j_1'', \ldots, j_{C''}''$ of $j_1, \ldots, j_C$.

Therefore, the set D contains at most

$$\ll 1^R \exp\left(\ell\sum_{n=1}^{L'}a_{i'_n} + 2\ell\sum_{n=1}^{L''}a_{i''_n} + \ell\sum_{n=1}^{C'}a_{j'_n} + bC' - \sum_{n=1}^{C'}n_{j'_n} + 2(\ell\sum_{n=1}^{C''}a_{j''_n} + bC'' - \sum_{n=1}^{C''}n_{j''_n})\right)$$

$$= \exp\left(\ell\left(\sum_{n=1}^{L'}a_{i'_n} + 2\sum_{n=1}^{L''}a_{i''_n} + \sum_{n=1}^{C'}a_{j'_n} + 2\sum_{n=1}^{C''}a_{j''_n}\right) + b(C' + 2C'') - \sum_{n=1}^{C'}n_{j'_n} - 2\sum_{n=1}^{C''}n_{j''_n}\right)$$

$$= \exp\left(\ell\left(\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \sum_{n=1}^{C}(a_{j_n}k_{j_n})\right) + b\sum_{n=1}^{C}k_{j_n} - \sum_{n=1}^{C}(n_{j_n}k_{j_n}))\right)$$

many disjoint sets of the form E.

On the other hand, Lemma 5.3 gives

$$(\ell+1)\left(\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \sum_{n=1}^{C}(a_{j_n}k_{j_n}) - \sum_{n=1}^{R}(a_{k_n}k_{k_n})\right) < 2\sum_{n=1}^{C}(s_{j_n}k_{j_n})$$

where $s_{j_k} \in J_{j_k}(V_m^{\mathcal{V}}) - b = (n_{j_k} - b, n_{j_k} + a_{j_k} - b]$. Thus,

$$(\ell+1)\left(\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \sum_{n=1}^{C}(a_{j_n}k_{j_n}) - \sum_{n=1}^{R}(a_{k_n}k_{k_n})\right) < 2\sum_{n=1}^{C}((n_{j_n} + a_{j_n} - b)k_{j_n})$$

and since $\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \sum_{n=1}^{C}(a_{j_n}k_{j_n}) + \sum_{n=1}^{R}(a_{k_n}k_{k_n}) = h_r + 2h_s = h_{\max}(\mathrm{T})$ we obtain

$$(\ell+1)\left(2\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + 2\sum_{n=1}^{C}(a_{j_n}k_{j_n}) - h_{\max}(\mathrm{T})\right) < 2\sum_{n=1}^{C}((n_{j_n} + a_{j_n} - b)k_{j_n}).$$

Dividing both sides by 2 and simplifying the expression we get

$$\ell\left(\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \sum_{n=1}^{C}(a_{j_n}k_{j_n})\right) + b\sum_{n=1}^{C}k_{j_n} - \sum_{n=1}^{C}(n_{j_n}k_{j_n})$$
$$< -\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \frac{(\ell+1)h_{\max}(\mathrm{T})}{2}.$$

Hence, the set D can be decomposed into

$$\ll \exp\left(-\sum_{n=1}^{L}(a_{i_n}k_{i_n}) + \frac{(\ell+1)h_{\max}(\mathrm{T})}{2}\right) \ll \exp\left(\frac{h_{\max}(\mathrm{T})}{2}\ell\right)$$

disjoint sets of the form E. $\qquad\square$

5.2. The proof of Proposition 5.1.
Let $P(\mathcal{V}) \in P_{M,N}$ be given. Since $X_{<M}$ is pre-compact it suffices to restrict ourselves to a neighborhood $\mathcal{O}$ of some $x_0 \in X_{<M} \cap P(\mathcal{V})$. We let $\mathcal{O} = x_0 B_{\eta/2}^{U^+}B_{\eta/2}^{U^-L}$ be a neighborhood of such $x_0 \in X_{<M}$ and define the set $P_{\mathcal{O}}(\mathcal{V})$ by

$$P_{\mathcal{O}}(\mathcal{V}) = \mathcal{O} \cap P(\mathcal{V}).$$

It suffices to prove that the set $P_{\mathcal{O}}(\mathcal{V})$ can be covered by $\ll c_0^{\frac{h_{\max}(\mathrm{T})N}{\log M}}e^{h_{\max}(\mathrm{T})(N-\frac{1}{2}|\mathcal{V}|)}$ Bowen N-balls for some universal constant $c_0 \geq 1$.

Let us make some observations. If we consider the image of $\mathcal{O}$ under T^n we obtain the set

$$\mathrm{T}^n(\mathcal{O}) = \mathrm{T}^n(x_0)(a^{-n}B_{\eta/2}^{U^+}a^n)a^{-n}B_{\eta/2}^{U^-}La^n.$$

We see that the jth component of the U^+-part gets stretched by the factor e^{na_j}. Here again we naturally embed U^+ into $\mathbb{R}^r \times \mathbb{C}^s$. Under this identification, dividing $(a^{-n}B_{\eta/2}^{U^+}a^n)$ into $\prod_{j=1}^{r+s}\lceil e^{na_j}\rceil^{\delta_j}$ many small parts we obtain the sets of the form

$$\mathrm{T}^n(x_0)u^+B_{\eta/2}^{U^+}a^{-n}B_{\eta/2}^{U^-}La^n$$

for some $u^+ \in U^+$. Now, if we take the pre-image under T^n of these sets then we obtain the similar sets

$$\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+)a^nB_{\eta/2}^{U^+}a^{-n}B_{\eta/2}^{U^-}L$$

as before. It is not hard to see that the set $\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+)a^nB_{\eta/2}^{U^+}a^{-n}B_{\eta/2}^{U^-}L$ is contained in the forward Bowen n-ball $\mathrm{T}^{-n}(\mathrm{T}^n(x_0)u^+)B_n^+$. This in particular shows that $\mathcal{O}$ can be covered by $\ll e^{(h_r+2h_s)n}$ many forward Bowen n-balls which is the reason why the maximal entropy $h_{\max}(\mathrm{T})$ is $h_r + 2h_s$. However, using Lemma 5.4 we will show that we in fact need fewer Bowen balls to cover the set $\mathcal{O}$.

Let us recall that we decompose $\mathcal{V}$ into ordered maximal subintervals $V_m^{\mathcal{V}}$ so that we have

$$\mathcal{V} = V_1^{\mathcal{V}} \cup V_2^{\mathcal{V}} \cup ... \cup V_k^{\mathcal{V}}.$$

Now we let $[0, N-1] \backslash \mathcal{V} = W_1 \cup W_2 \cup ... \cup W_{k'}$ where W_m are again ordered maximal intervals. We inductively prove the following:

If $[0, b-1] = V_1^{\mathcal{V}} \cup V_2^{\mathcal{V}} \cup ... \cup V_{m-1}^{\mathcal{V}} \cup W_1 \cup W_2 \cup ... \cup W_{n'}$ then for some constant c_0 the set $P_{\mathcal{O}}(\mathcal{V})$ can be covered by

$$\leq c_0^{m-1+n'} \exp\left(h_{\max}(\mathrm{T})\left[(b-1) - \frac{(|V_1^{\mathcal{V}}| + \cdots + |V_{m-1}^{\mathcal{V}}|)}{2}\right]\right)$$

pre-images under T^{b-1} of sets of the form

(5.6) $$\mathrm{T}^{b-1}(x_0)u^+B_{\eta/2}^{U^+}a^{-b+1}B_{\eta/2}^{U^-}La^{b-1}.$$

For the interval $[0,0]$ the claim is obvious. Now, assume that the claim is true for the interval $[0, b-1]$ as above. In the inductive step, if the next interval is $W_{n'+1}$ then once we divide each set obtained earlier into

$$\prod_{j=1}^{r+s}\lceil e^{|W_{n'+1}|a_j}\rceil^{\delta_j} \leq c_0 e^{h_{\max}(\mathrm{T})(|W_{n'+1}|)}$$

small ones for some constant c_0, we just keep all of them. So, assume that the next interval is $V_m^{\mathcal{V}} = [b, b+\ell]$. Let Y be one of the sets (5.6) obtained in the earlier step. We would like to estimate the upper bound to cover Y by pre-images under T^{ℓ} of sets of the form

(5.7) $$\mathrm{T}^{b-1+\ell}(x_0)u^+(\mathbf{t})B_{\eta/2}^{U^+}a^{-b+1-\ell}B_{\eta/2}^{U^-}La^{b-1+\ell}.$$

We are interested in the points $x \in Y$ for which $\mathrm{T}^{-b+1}(x)$ is in $Q(J(V_m^{\mathcal{V}}))$. We know by assumption that x_0 is one of them. If $x \in Y$ is another one then by Lemma 5.2 there exists $\mathbf{t} \in B_{\eta/2}^{\mathbb{R}^r \times \mathbb{C}^s}$ such that $x = x_0 u^+(\mathbf{t})g$ for some $g \in B_{\eta/2}^{U^-}L$ and for $j \in [1, r+s]$, $|t_j| \ll e^{b-n_j}$ where n_j is the left end point of the interval $J_j(V_m^{\mathcal{V}})$. Hence the set we are interested in corresponds to the set D in Lemma 5.4

and each set as in (5.7) corresponds to the set E as in Lemma 5.4. Thus, if necessary enlarging the constant c_0 appeared earlier, using Lemma 5.4 we see that once we divide Y into the sets of the form as in (5.7) we only need to keep

$$\leq c_0 e^{\frac{h_{\max}(\mathrm{T})}{2}\ell} = c_0 e^{h_{\max}(\mathrm{T})(\ell - \frac{|V_m^{\mathcal{V}}|}{2})}$$

many of them. Hence, we conclude that the set $P_{\mathcal{O}}(\mathcal{V})$ can be covered by

$$\leq c_0^{m+n'} \exp\left(h_{\max}(\mathrm{T})\left[(b-1+\ell) - \frac{(|V_1^{\mathcal{V}}| + \cdots + |V_m^{\mathcal{V}}|)}{2}\right]\right)$$

pre-images under $\mathrm{T}^{b+\ell-1}$ of the sets of the form

$$\mathrm{T}^{b+\ell-1}(x_0)u^+(t)B_{\eta/2}^{U^+}a^{-b-\ell+1}B_{\eta/2}^{U^-L}a^{b+\ell-1}.$$

Now, we let $b = N$ to obtain that the set $P_{\mathcal{O}}(\mathcal{V})$ can be covered by

$$\leq c_0^{k+k'} \exp\left(h_{\max}(\mathrm{T})\left[N - \frac{(|V_1^{\mathcal{V}}| + \cdots + |V_k^{\mathcal{V}}|)}{2}\right]\right)$$

many Bowen N-balls. On the other hand, the proof of Lemma 3.2 suggests that k and hence k' are bounded above by

$$\frac{N}{2\lfloor \frac{2\log M}{h_r + h_s}\rfloor} + 1.$$

Thus, the set $P_{\mathcal{O}}(\mathcal{V})$ can be covered by

$$\ll c_0^{\frac{h_{\max}(\mathrm{T})N}{\log M}} e^{h_{\max}(\mathrm{T})(N - \frac{|\mathcal{V}|}{2})}$$

translates of Bowen N-balls, which completes the proof.   □

6. Proof of Theorem 1.4

We recall that $a_* = \max\{a_1, a_2, \ldots, a_{r+s}\}$ and $D = \dim U^+$. One can easily obtain the following lemma.

Lemma 6.1. *For any $N \in \mathbb{N}$, the set $a^N B_{\eta/2}^{U^+} a^{-N} B_{\eta/2}^{U^-L}$ can be covered by $\ll e^{[Da_* - h_{\max}(\mathrm{T})]N}$ translates of $B_{\frac{\eta}{2}e^{-a_*N}}^{U^+} B_{\frac{\eta}{2}}^{U^-L}$.*

The proof of Proposition 5.1 together with Lemma 6.1 at once give

Proposition 6.2. *The set $P(\mathcal{V})$ can be covered by*

$$\ll_M c_0^{\frac{h_{\max}(\mathrm{T})}{\log M}N} e^{(DNa_* - \frac{h_{\max}(\mathrm{T})|\mathcal{V}|}{2})}$$

translates of $B_{\frac{\eta}{2}e^{-a_*N}}^{U^+} B_{\frac{\eta}{2}}^{U^-L}$ in X for some universal constant $c_0 \geq 1$.

Proof. From the proof of Proposition 5.1 we know that the set $P(\mathcal{V})$ can be covered by

$$\ll_M c_0^{\frac{h_{\max}(\mathrm{T})}{\log M}N} e^{h_{\max}(\mathrm{T})(N - \frac{|\mathcal{V}|}{2})}$$

translates of $a^N B_{\eta/2}^{U^+} a^{-N} B_{\eta/2}^{U^-L}$ in X. Thus, Lemma 6.1 finishes the proof.   □

For any fixed $N \in \mathbb{N}$ and for any $x \in X$ we associate $\mathcal{V}_x \subset [0, N-1]$ such that for any $n \in [0, N-1]$

$$\mathrm{T}^n(x) \in X_{\geq M} \text{ if and only if } n \in \mathcal{V}_x.$$

As in the introduction, let ν be a measure on X of dimension d in the unstable direction. For any $\delta > 0$ we note that $\nu(B^{U^+}_{\frac{\eta}{2}e^{-a_*N}} B^{U^- L}_{\frac{\eta}{2}}) \ll e^{-a_*(d-\delta)N}$. Using Proposition 5.1 together with Lemma 4.1 and Lemma 3.2 it is easy to obtain the following (cf. [2, Lemma 6.2]).

Lemma 6.3. *For any $\delta > 0, \kappa \in [0,1]$ and for any $N, M \geq 1$ large, we have*

$$\nu(x \in X_{<M} : |\mathcal{V}_x| > \kappa N)$$

$$\ll_{M,\delta} \exp\left(\left[Da_* - \frac{h_{\max}(\mathrm{T})}{2}\kappa - (d-\delta)a_* + O\left(\frac{\log \log M}{\log M}\right)\right] N\right).$$

Proof of Theorem 1.4. We follow the proof of [2, Theorem 1.6]. The conclusion of the theorem is trivial when $d \leq D - \frac{h_{\max}(\mathrm{T})}{2a_*}$ so that we may assume $d > D - \frac{h_{\max}(\mathrm{T})}{2a_*}$. We first estimate an upper bound for $\mu_N(X_{\geq M})$ when $M, N \geq 1$ large. We have

$$\mu_N(X_{\geq M}) = \frac{1}{N} \sum_{n=0}^{N-1} \nu(\mathrm{T}^{-n}(X_{\geq M}))$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{<M} \cap \mathrm{T}^{-n}(X_{\geq M})) + \frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{>M} \cap \mathrm{T}^{-n}(X_{\geq M}))$$

$$\leq \frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{<M} \cap \mathrm{T}^{-n}(X_{\geq M})) + \nu(X_{\geq M}).$$

It suffices to estimate is $\frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{<M} \cap \mathrm{T}^{-n}(X_{\geq M}))$. For this, we note that

$$\frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{\leq M} \cap \mathrm{T}^{-n}(X_{\geq M}))$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{W \subset [0,N]} \nu(\{x \in X_{<M} : V_x = W\} \cap \mathrm{T}^{-n}(X_{\geq M})),$$

where the term $\nu(\{x \in X_{<M} : V_x = W\} \cap \mathrm{T}^{-n}(X_{\geq M}))$ is either 0 or is equal to $\nu(\{x \in X_{<M} : V_x = W\})$. Switching the order of summation yields

$$= \frac{1}{N} \sum_{W \subset [0,N-1]} |W| \nu(\{x \in X_{<M} : V_x = W\})$$

$$= \frac{1}{N} \sum_{i=1}^{N} i \nu(\{x \in X_{<M} : |V_x| = i\})$$

$$= \frac{1}{N} \sum_{i=1}^{\lfloor \kappa N \rfloor} i \nu(\{x \in X_{<M} : |V_x| = i\}) + \frac{1}{N} \sum_{i=\lceil \kappa N \rceil}^{N} i \nu(\{x \in X_{<M} : |V_x| = i\})$$

$$\leq \frac{1}{N} \lfloor \kappa N \rfloor \nu(X_{<M}) + \frac{1}{N} N \nu(\{x \in X_{<M} : |V_x| > \kappa N\})$$

Let $K(M, \delta) > 0$ be the implicit constant appeared in Lemma 6.3. Then we obtain

$$\frac{1}{N} \sum_{n=0}^{N-1} \nu(X_{<M} \cap \mathrm{T}^{-n}(X_{\geq M})) \leq \kappa + K(M, \delta) e^{\left(Da_* - \frac{h_{\max}(\mathrm{T})}{2}\kappa - (d-\delta)a_* + O\left(\frac{\log \log M}{\log M}\right)\right)N}.$$

Therefore we get that

$$(6.1) \quad \mu_N(X_{\geq M}) \leq \epsilon(M) + \kappa + K(M, \delta) e^{\left(Da_* - \frac{h_{\max}(\mathrm{T})}{2}\kappa - (d-\delta)a_* + O\left(\frac{\log \log M}{\log M}\right)\right)N}.$$

By assumption we have $d > D - \frac{h_{\max}(\mathrm{T})}{2a_*}$ so that $2a_*(D - d)/(h_{\max}(\mathrm{T})) < 1$. Now, for any $\kappa \in (2a_*(D-d)/(h_{\max}(\mathrm{T})), 1]$ we may pick $\delta > 0$ small enough so that

$$Da_* - \frac{h_{\max}(\mathrm{T})}{2}\kappa - (d-\delta)a_* + O\left(\frac{\log \log M}{\log M}\right) < 0$$

for sufficiently large M. Thus, for any $\epsilon > 0$ we may choose M sufficiently large so that

$$\mu_N(X_{\geq M}) \leq \kappa + \epsilon$$

which gives in the limit that $\mu(X) > 1 - \kappa$. This holds for any $\kappa > 2a_*(D - d)/(h_{\max}(\mathrm{T}))$. Thus,

$$\mu(X) \geq 1 - \frac{2a_*(D-d)}{h_{\max}(\mathrm{T})}.$$

$\square$

REFERENCES

[1] S. G. Dani. On orbits of unipotent fows on homogeneous spaces. *Ergodic Theory Dynam. Systems*, 6:167–182, 1986.

[2] M. Einsiedler and S. Kadyrov. Escape of mass and entropy for $\mathrm{SL}_3(\mathbb{Z})\backslash \mathrm{SL}_3(\mathbb{R})$. *to appear in Israel J. Math.*, 2011.

[3] M. Einsiedler, E. Lindenstrauss, Ph. Michel, and A. Venkatesh. Distribution of closed geodesics on the modular surface, and Duke's theorem. *Preprint*, 2011.

[4] M. Einsiedler and T. Ward. *Ergodic Theory with a view towards Number Theory*. Springer, 2011.

[5] K. Falconer. *Fractal Geometry*. Wiley, 2003.

[6] S. Kadyrov. Positive entropy invariant measures on the space of lattices with escape of mass. *to appear in Ergodic Theory Dynam. Systems*, 2011.

[7] G. A. Margulis and D. Kelinbock. Flows on homogeneous spaces and diophantine approximation on manifolds. *Ann. Math.*, 148:339–360, 1998.

[8] S. Mozes and N. Shah. On the space of ergodic invariant measures of unipotent flows. *Ergodic Theory Dynam. Systems*, 15:149–159, 1995.

[9] R. Shi. Equidistribution of expanding measures with local maximal dimension and diophantine approximation. *arXiv:0905.1152v2*, 2009.

[10] A. Sleberg. Recent developments in the theory of discontinuous groups of motions of symmetric spaces. *1970 Proceedings of the Fifteenth Scandinavian Congress (Oslo, 1968) Lecture Notes in Mathematics, Springer*, 118:99–120, 1970.

[11] Peter Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, 2000.